

WHAT IS CLAIMED IS:

1. A digital watermark embedding apparatus  
comprising:

5 a first generation unit configured to generate a  
plurality of symbol sequences each of which includes a  
plurality of symbols including ranks, each of the ranks  
being uniquely numbered among each of the symbol  
sequences, each of the symbol sequences uniquely  
corresponding to each of a plurality of identification  
10 information items to be embedded as digital watermark  
information into each of copies of digital contents;

a second generation unit configured to generate a  
plurality of to-be-embedded codes corresponding to each  
of the symbols in each of the symbol sequences; and

15 an embedding unit configured to embed the to-be-  
embedded codes in each of the copies.

2. The digital watermark embedding apparatus  
according to claim 1, wherein the second generation  
unit generates the to-be-embedded codes in a random  
20 number sequence such that the to-be-embedded codes  
corresponding to symbols including a same rank in the  
symbol sequences include one of no cross-correlation  
and a very low cross-correlation.

3. The digital watermark embedding apparatus  
25 according to claim 2, wherein the second generation  
unit generates the to-be-embedded codes in a random  
number sequence such that the to-be-embedded codes

corresponding to symbols differing in rank in the symbol sequences include one of no cross-correlation and a very low cross-correlation.

4. The digital watermark embedding apparatus  
5 according to claim 1, wherein:

the identification information items include a plurality of integers;

a plurality of particular integers are preset for each of the ranks of each of the symbol sequences; and

10 the first generation unit divides each of the identification information items by each of the particular integers to obtain a plurality of residuals, arranges, in accordance with the ranks, the residuals corresponding to each of the ranks, and outputs the  
15 residuals arranged as an output symbol sequence.

5. The digital watermark embedding apparatus according to claim 4, wherein the particular integers are relatively prime to each other.

6. The digital watermark embedding apparatus  
20 according to claim 1, wherein:

the identification information items include a plurality of integers;

a plurality of particular integers are preset for each of the ranks of each of the symbol sequences; and

25 the first generation unit divides each of the identification information items by each of the particular integers to obtain a plurality of residuals,

converts, into a plurality of unique information items,  
the residuals corresponding to each of the ranks,  
arranges the unique information items in accordance  
with the ranks, and outputs the unique information  
5 items arranged as an output symbol sequence.

7. The digital watermark embedding apparatus  
according to claim 6, wherein the particular integers  
are relatively prime to each other.

8. The digital watermark embedding apparatus  
10 according to claim 1, wherein a range of an  
identification information amount which can be embedded  
as watermark information into the copies is set  
narrower than a range of an identification information  
amount to which the symbol sequence is assigned.

15 9. The digital watermark embedding apparatus  
according to claim 1, wherein the identification  
information items include unique information assigned  
to users who provide the copies.

10. A digital watermark analysis apparatus for  
20 specifying at least one of a plurality of identifica-  
tion information items embedded as a plurality of  
watermark information items in a plurality of legal  
copies of digital contents used for collusive attacks,  
from a plurality of illegal copies of the digital  
25 contents obtained by collusive attacks made against the  
legal copies, the digital watermark analysis apparatus  
comprising:

an extraction unit configured to extract a plurality of embedded codes including ranks from the illegal copies, each of the ranks being uniquely numbered among each of the symbol sequences;

5           an acquisition unit configured to acquire a plurality of symbols corresponding to the embedded codes and arrange the symbols in accordance with the ranks of the embedded codes, and acquire a first symbol sequence of symbol sequences each of which includes a plurality of the symbols based on the symbols; and

10           a specifying unit configured to specify at least one of the identification information items embedded in the legal copies, based on the first symbol sequence and second symbol sequences uniquely assigned to the identification information items.

11. The digital watermark analysis apparatus according to claim 10, wherein:

20           the specifying unit specifies at least one of the identification information items based on the second symbol sequences each of which include a plurality of symbols including the ranks; and

25           the acquisition unit acquires a plurality of the symbols including the ranks, the symbols corresponding to the embedded codes, the embedded codes being embedded in the legal copies.

12. The digital watermark analysis apparatus according to claim 10, wherein the extraction unit

extracts, in units of the ranks, one of the embedded codes corresponding to the symbols, the one of the embedded codes being estimated to be most frequently used in the collusive attacks.

5           13. The digital watermark analysis apparatus according to claim 10, wherein:

the extraction unit extracts several of the embedded codes which are generated in a random number sequence such that the embedded codes corresponding to  
10 the symbols including a same rank in the symbol sequences include one of no cross-correlation and a very low cross-correlation; and

the extraction unit obtains, in units of the ranks, cross-correlation between embedded codes  
15 corresponding to the symbols contained in the first symbol sequence, and obtains one of the embedded codes which includes a maximum cross-correlation.

14. The digital watermark analysis apparatus according to claim 13, wherein the extraction unit  
20 extracts several of the embedded codes which are generated in a random number sequence such that the embedded codes corresponding to the symbols differing in each of the ranks in the symbol sequences include one of no cross-correlation and a very low cross-  
25 correlation.

15. The digital watermark analysis apparatus according to claim 10, wherein the acquisition unit

compares the second symbol sequences with third symbol sequences extracted from the illegal copies, and specifies at least one of the identification information items embedded in the legal copies based on  
5 comparison results of the acquisition unit.

16. The digital watermark analysis apparatus according to claim 15, wherein the acquisition unit compares, in units of the ranks, the second symbol sequences with the third symbol sequences, and if the  
10 third symbol sequences are identical at the ranks greater than a preset number of the ranks, the identification information items are determined to be those embedded in the legal copies.

17. The digital watermark analysis apparatus  
15 according to claim 16, wherein

if the identification information items are a plurality of integers, a plurality of particular integers  $N(i)$  ( $i$ : integer,  $N(1) \leq N(2) \leq \dots \leq N(M)$ ) are preset for each of the ranks  $i$  of each of the second  
20 symbol sequences, and the extraction unit selects, from a range of  $0 \sim N(i)-1$ , a symbol  $S(i)$  corresponding to each of the ranks  $i$  of the second symbol sequences, and arranges the symbol  $S(i)$ , into a to-be-generated symbol sequence, in accordance with each of the ranks  $i$ ,  
25 the preset number is  $(k + 1)$ ,  
where  $k$  is a value which makes a product of  $N(1), \dots, N(k)$  higher than a total number of the identification

information items, and  $\underline{l}$  satisfies  $[1 - \prod_{i=1}^l 1/N(i)]^S \geq 1 - \varepsilon_2$  (a range of  $i$  that assumes  $\Pi$  is  $i = 1 \sim l$  or  $i = k + 1 \sim (k + l)$ ,  $S = M^{C_{k+1}}$ , and  $\varepsilon_2$  ( $0 < \varepsilon_2 < 1$ ) represents a rate of error tracing at which erroneous identification information is specified as the identification information embedded in the legal copies).

18. The digital watermark analysis apparatus according to claim 16, wherein

10        if the identification information items are a plurality of integers, a plurality of particular integers  $N(i)$  ( $\underline{i}$ : integer,  $q = N(1) = N(2) = \dots = N(M)$ ) are preset for each of the ranks  $\underline{i}$  of each of the second symbol sequences, and the extraction unit  
15        selects, from a range of  $0 \sim N(i)-1$ , a symbol  $S(i)$  corresponding to each of the ranks  $\underline{i}$  of the second symbol sequences, and arranges the symbol  $S(i)$ , into a to-be-generated symbol sequence, in accordance with each of the ranks  $\underline{i}$ ,

20        the predetermined number is  $(k + l)$ , where  $\underline{k}$  is a value which makes a product of  $N(1), \dots, N(k)$  higher than a total number of the identification information items, and  $\underline{l}$  satisfies  $[1 - 1/q^l]^S \geq 1 - \varepsilon$ , where  $S = M^{C_{k+1}}$ , and  $\varepsilon$  ( $0 < \varepsilon < 1$ ) represents a  
25        rate of error tracing at which erroneous identification information is specified as the identification information embedded in the legal copies.

19. The digital watermark analysis apparatus according to claim 10, wherein a maximum number of the legal copies is set to a preset value, and the acquisition unit specifies at least one of collusion  
5 groups which form the symbol sequences extracted from the legal copies, and determines at least one of the identification information items embedded in the legal copies based on at least one collusion group, the collusion groups being each formed of a combination of  
10 identification information items assigned to a number of legal copies not larger than the maximum number of legal copies.

20. The digital watermark analysis apparatus according to claim 19, wherein if only one of the  
15 collusion groups is specified, the acquisition unit specifies all identification information items forming the one collusion group, as the identification information items embedded in the legal copies.

21. The digital watermark analysis apparatus  
20 according to claim 19, wherein if a plurality of groups not less than two collusion groups are specified, the acquisition unit specifies only common identification information of the collusion groups as identification information embedded in the legal copies.

22. The digital watermark analysis apparatus  
25 according to claim 10, wherein a range of an identification information amount which is embedded as



watermark information into the legal copies is set narrower than a range of an identification information amount to which the symbol sequences is assigned.

23. The digital watermark analysis apparatus  
5 according to claim 10, wherein the identification information items includes unique information assigned to users who provide the legal copies.

24. A digital watermark embedding method comprising:

10       generating a plurality of symbol sequences each of which includes a plurality of symbols including ranks, each of the ranks being uniquely numbered among each of the symbol sequences, each of the symbol sequences uniquely corresponding to each of a plurality of  
15 identification information items to be embedded as digital watermark information into each of copies of digital contents;

      generating a plurality of to-be-embedded codes corresponding to each of the symbols in each of the  
20 symbol sequences; and

      embedding the to-be-embedded codes in each of the copies.

25. A digital watermark analysis method of specifying at least one of a plurality of identifica-  
25 tion information items embedded as a plurality of watermark information items in a plurality of legal copies of digital contents used for collusive attacks,

from a plurality of illegal copies of the digital contents obtained by collusive attacks made against the legal copies, the digital watermark analysis method comprising:

5           extracting a plurality of embedded codes including ranks from the illegal copies, each of the ranks being uniquely numbered among each of the symbol sequences;

          acquiring a plurality of symbols corresponding to the embedded codes and arranging the symbols in  
10       accordance with the ranks of the embedded codes, and acquiring a first symbol sequence of symbol sequences each of which includes a plurality of the symbols based on the symbols; and

          specifying at least one of the identification  
15       information items embedded in the legal copies, based on the first symbol sequence and second symbol sequences uniquely assigned to the identification information items.

26. A program stored in a computer readable  
20       medium, comprising:

          means for instructing a computer to generate a plurality of symbol sequences each of which includes a plurality of symbols including ranks, each of the ranks being uniquely numbered among each of the symbol  
25       sequences, each of the symbol sequences uniquely corresponding to each of a plurality of identification information items to be embedded as digital watermark

information into each of copies of digital contents;

means for instructing the computer to generate a plurality of to-be-embedded codes corresponding to each of the symbols in each of the symbol sequences; and

5 means for instructing the computer to embed the to-be-embedded codes in each of the copies.

27. A program stored in a computer readable medium which enables a computer to function as a digital watermark analysis apparatus for specifying at least one of a plurality of identification information items  
10 embedded as a plurality of watermark information items in a plurality of legal copies of digital contents used for collusive attacks, from a plurality of illegal copies of the digital contents obtained by collusive  
15 attacks made against the legal copies, the program comprising:

means for instructing the computer to extract a plurality of embedded codes including ranks from the illegal copies, each of the ranks being uniquely  
20 numbered among each of the symbol sequences;

means for instructing the computer to acquire a plurality of symbols corresponding to the embedded codes and arranging the symbols in accordance with the ranks of the embedded codes, and acquiring a first  
25 symbol sequence of symbol sequences each of which includes a plurality of the symbols based on the symbols; and

means for instructing the computer to specify at  
least one of the identification information items  
embedded in the legal copies, based on the first symbol  
sequence and second symbol sequences uniquely assigned  
5 to the identification information items.